

Эти инструменты конкурируют с моделями Sora и Dall-E 3 от OpenAI. Google утверждает, что Veo обладает «продвинутым пониманием естественного языка и визуальной семантики», позволяя создавать видео на основе текста.

Veo может создавать видео более минуты и понимает такие техники, как таймлапс. В промо-ролике Дональд Гловер и его студия Gilga демонстрируют, как с помощью текста создается видео с кабриолетом и парусником. Veo улучшил симуляцию физики и качество видео.

Veo будет доступен в инструменте VideoFX от Google и появится в YouTube Shorts.

Что касается Imagen 3, Google обещает «высочайшее качество» модели для преобразования текста в изображение с меньшим количеством артефактов. Эта модель лучше справляется с длинными текстовыми запросами и сложными деталями.